PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N

A2

(11) International Publication Number: WO 99/54437

(43) International Publication Date: 28 October 1999 (28.10.99)

(21) International Application Number: PCT/US99/08900

(22) International Filing Date: 23 April 1999 (23.04.99)

(30) Priority Data:

09/065,363	23 April 1998 (23.04.98)	US
Not furnished	23 April 1999 (23.04.99)	US

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

Without international search report and to be republished upon receipt of that report.

(71) Applicant: MILLENNIUM BIOTHERAPEUTICS, INC. [US/US]; 620 Memorial Drive, Cambridge, MA 02139 (US).

(72) Inventor: HOLTZMAN, Douglas, A.; 821 Centre Street #6, Jamaica Plain, MA 01230 (US).

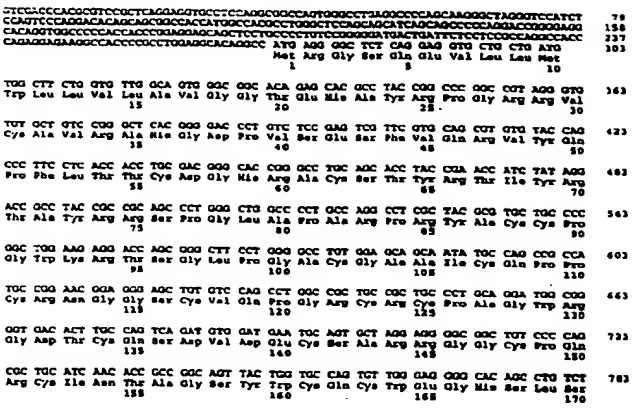
(74) Agent: MEIKLEJOHN, Anita, L.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).

(54) Title: NOVEL MOLECULES OF THE T125-RELATED PROTEIN FAMILY AND USES THEREOF

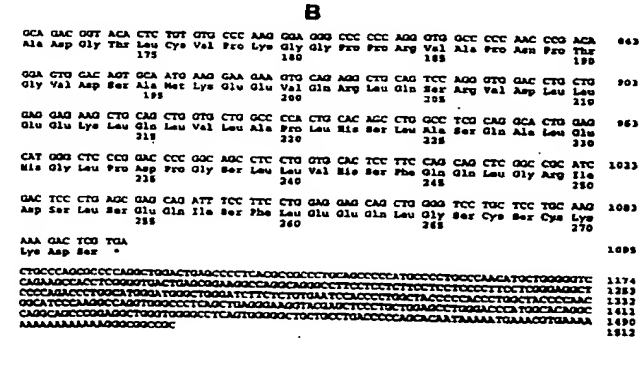
(57) Abstract

Novel T125 polypeptides, proteins, and nucleic acid molecules are disclosed. In addition to isolated, full-length T125 proteins, the invention further provides isolated T125 fusion proteins, antigenic peptides and anti-T125 antibodies. The invention also provides T125 nucleic acid molecules, recombinant expression vectors containing a nucleic acid molecule of the invention, host cells into which the expression vectors have been introduced and non-human transgenic animals in which a T125 gene has been introduced or disrupted. Diagnostic, screening and therapeutic methods utilizing compositions of the invention are also provided.

A



B



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

NOVEL MOLECULES OF THE T125-RELATED PROTEIN FAMILY AND USES THEREOF

Background of the Invention

The invention relates to a novel secreted protein and the gene encoding it.

Many membrane-associated and secreted proteins, for example, cytokines, play a vital role in the regulation of cell growth, cell differentiation, and a variety of specific cellular responses. A number of medically useful proteins, including erythropoietin, granulocyte-macrophage colony stimulating factor, human growth hormone, and various interleukins, are secreted proteins. Thus, an important goal in the design and development of new therapies is the identification and characterization of membrane-associated and secreted proteins and the genes which encode them.

Many membrane-associated proteins are receptors which bind a ligand and transduce an intracellular signal, leading to a variety of cellular responses. The identification and characterization of such a receptor enables one to identify both the ligands which bind to the receptor and the intracellular molecules and signal transduction pathways associated with the receptor, permitting one to identify or design modulators of receptor activity, e.g., receptor agonists or antagonists and modulators of signal transduction.

Summary of the Invention

The present invention is based, at least in part, on the discovery of a gene encoding T125. The T125 cDNA described below (SEQ ID NO:1) has a 819 nucleotide open reading frame (nucleotides 274-1092 of SEQ ID NO:1; SEQ ID NO:3) which encodes a 273 amino acid protein (SEQ ID NO:2). This protein includes a predicted signal sequence of about 22 amino acids (from amino acid 1 to about amino acid 22 of SEQ ID NO:2) and a predicted mature protein of about 252 amino acids (from about amino acid 23 to amino acid 274 of SEQ ID NO:2; SEQ ID NO:4). T125 protein possesses two epidermal growth factor (EGF)-like domains: amino acids 107 to 134 (SEQ ID NO:5), referred to herein as the "EGF1 domain", and amino acids 141 to 176 (SEQ ID NO:6), referred to herein as the "EGF2 domain". T125 is predicted to have no transmembrane domains and appears to be a secreted protein.

In addition, there are three additional alternatively spliced forms of human T125: T125a, T125b, and T125c. Figure 5 depicts the cDNA sequence (SEQ ID NO:11) and predicted amino acid sequence (SEQ ID NO:12) of human T125a. The open reading frame of SEQ ID NO:11 extends from nucleotide 94 to nucleotide 442 of SEQ ID NO:11 (SEQ ID NO:13). Figure 6 depicts the cDNA sequence (SEQ ID NO:14) and predicted amino acid sequence (SEQ ID NO:15) of human T125b. The open reading frame of SEQ ID NO:14 extends from nucleotide 194 to nucleotide 934 of SEQ ID NO:14 (SEQ ID NO:16). Figure 7 depicts the cDNA sequence (SEQ ID NO:17) and predicted amino acid sequence (SEQ ID NO:18) of T125c. The open reading frame of SEQ ID NO:17 extends from nucleotide 194 to nucleotide 823 of SEQ ID NO:17 (SEQ ID NO:19).

Unless otherwise specified, "T125" (or "TANGO 125") is used to refer to all forms of T125 (T125, T125a, T125b, and T125c).

The T125 molecules of the present invention are useful as modulating agents in regulating a variety of cellular processes. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding T125 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of T125-encoding nucleic acids.

The invention features a nucleic acid molecule which is at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the nucleotide sequence shown in SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693 (the "cDNA of ATCC 98693"), or a complement thereof.

The invention features a nucleic acid molecule which includes a fragment of at least 425 (450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1290) nucleotides of the nucleotide sequence shown in SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the nucleotide sequence of the cDNA of ATCC 98693, or a complement thereof.

The invention also features a nucleic acid molecule which includes a nucleotide sequence encoding a protein having an amino acid sequence that is at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the amino acid sequence of SEQ ID NO:2, 4, 12, 15, or 18, or the amino acid sequence encoded by the cDNA of ATCC 98693. In a preferred embodiment, a T125 nucleic acid molecule has the nucleotide sequence shown SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the nucleotide sequence of the cDNA of ATCC 98693.

Also within the invention is a nucleic acid molecule which encodes a fragment of a polypeptide having the amino acid sequence of SEQ ID NO:2, 4, 12, 15, or 18, the fragment including at least 15 (25, 30, 50, 100, 150,

300, or 400) contiguous amino acids of SEQ ID NO:2, 4, 12, 15, or 18, or the polypeptide encoded by the cDNA of ATCC Accession Number 98693.

The invention includes a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 4, 12, 15, or 18, or an amino acid sequence encoded by the cDNA of ATCC Accession Number 98693, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, under stringent conditions.

Also within the invention are: an isolated T125 protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:4 (mature human T125) or the amino acid sequence of SEQ ID NO:2 (immature human T125), SEQ ID NO:12 (human T125a), SEQ ID NO:15 (human T125b), or SEQ ID NO:18 (human T125c); and an isolated T125 protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the EGF1 or EGF2 domains of SEQ ID NO:2 (e.g., about amino acid residues 107 to 134 or 141 to 176 of SEQ ID NO:2).

Also within the invention are: an isolated T125 protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:3 or the cDNA of ATCC 98693; an isolated T125 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 65% preferably 75%, 85%, or 95% identical the EGF-like domain encoding portions of SEQ ID NO:1 (e.g., about nucleotides 592 to 675 or 694 to 801 of SEQ ID NO:1); and an isolated T125 protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the

nucleotide sequence of SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the non-coding strand of the cDNA of ATCC 98693.

Also within the invention is a polypeptide which is a naturally occurring allelic variant of a polypeptide that includes the amino acid sequence of SEQ ID NO:2, 4, 12, 15, or 18, or an amino acid sequence encoded by the cDNA insert of ATCC as Accession Number 98693, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, under stringent conditions.

Another embodiment of the invention features T125 nucleic acid molecules which specifically detect T125 nucleic acid molecules. For example, in one embodiment, a T125 nucleic acid molecule hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the cDNA of ATCC 98693, or a complement thereof. In another embodiment, the T125 nucleic acid molecule is at least 450 (500, 550, 600, 650, 700, 800, 900, 1000, or 1290) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, the cDNA of ATCC 98693, or a complement thereof. In a preferred embodiment, an isolated T125 nucleic acid molecule comprises nucleotides 592 to 675 or 694 to 801 of SEQ ID NO:1, encoding the EGF-like domains of T125, or a complement thereof. In another embodiment, the invention provides an isolated nucleic acid molecule which is antisense to the coding strand of a T125 nucleic acid.

Another aspect of the invention provides a vector, e.g., a recombinant expression vector, comprising a T125 nucleic acid molecule of the invention. In another

embodiment the invention provides a host cell containing such a vector. The invention also provides a method for producing T125 protein by culturing, in a suitable medium, a host cell of the invention containing a recombinant expression vector such that a T125 protein is produced.

Another aspect of this invention features isolated or recombinant T125 proteins and polypeptides. Preferred T125 proteins and polypeptides possess at least one biological activity possessed by naturally occurring human T125, e.g., (1) the ability to form protein:protein interactions with other proteins; (2) modulation of cellular proliferation; and (3) modulation of cellular differentiation.

The T125 proteins of the present invention, or biologically active portions thereof, can be operatively linked to a non-T125 polypeptide (e.g., heterologous amino acid sequences) to form T125 fusion proteins. The invention further features antibodies that specifically bind T125 proteins, such as monoclonal or polyclonal antibodies. In addition, the T125 proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of T125 activity or expression in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of T125 activity such that the presence of T125 activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating T125 activity comprising contacting a cell with an agent that modulates (inhibits or stimulates) T125 activity or expression such that T125 activity or expression in the cell is modulated. In one embodiment,

the agent is an antibody that specifically binds to T125 protein. In another embodiment, the agent modulates expression of T125 by modulating transcription of a T125 gene, splicing of a T125 mRNA, or translation of a T125 mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of the T125 mRNA or the T125 gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant T125 protein or nucleic acid expression or activity by administering an agent which is a T125 modulator to the subject. In one embodiment, the T125 modulator is a T125 protein. In another embodiment the T125 modulator is a T125 nucleic acid molecule. In other embodiments, the T125 modulator is a peptide, peptidomimetic, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant T125 protein or nucleic acid expression is a proliferative or differentiative disorder. T125 nucleic acid, polypeptides, and modulators of T125 expression or activity can be used to modulate wound healing and angiogenesis.

The present invention also provides a diagnostic assay for identifying the presence or absence of a genetic lesion or mutation characterized by at least one of: (i) aberrant modification or mutation of a gene encoding a T125 protein; (ii) mis-regulation of a gene encoding a T125 protein; and (iii) aberrant post-translational modification of a T125 protein, wherein a wild-type form of the gene encodes a protein with a T125 activity.

In another aspect, the invention provides a method for identifying a compound that binds to or modulates the activity of a T125 protein. In general,

such methods entail measuring a biological activity of a T125 protein in the presence and absence of a test compound and identifying those compounds which alter the activity of the T125 protein.

The invention also features methods for identifying a compound which modulates the expression of T125 by measuring the expression of T125 in the presence and absence of a compound.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1 depicts the cDNA sequence (SEQ ID NO:1) and predicted amino acid sequence (SEQ ID NO:2) of human T125 (also referred to as "TANGO 125"). The open reading frame of SEQ ID NO:1 extends from nucleotide 274 to nucleotide 1092 (SEQ ID NO:3).

Figure 2 depicts an alignment of the T125 amino acids 107 to 134 (EFG1) and 141 to 176 (EGF2) of SEQ ID NO:2 with the EGF-like domain consensus sequence derived from a hidden Markov model (PF00008; SEQ ID NO:7).

Figure 3 is a hydropathy plot of T125. The position of cysteines (cys) are indicated by the vertical bars immediately below the plot. Relative hydrophobicity is shown above the dotted line, and relative hydrophilicity is shown below the line.

Figure 4 depicts the cDNA sequence (SEQ ID NO:8) and predicted amino acid sequence (SEQ ID NO:10) of murine T125. The open reading frame of SEQ ID NO:8 extends from nucleotide 13 to nucleotide 837 of SEQ ID NO:8 (SEQ ID NO:9).

Figure 5 depicts the cDNA sequence (SEQ ID NO:11) and predicted amino acid sequence (SEQ ID NO:12) of human T125a, an alternatively spliced form of human T125. The open reading frame of SEQ ID NO:11 extends from

nucleotide 94 to nucleotide 442 of SEQ ID NO:11 (SEQ ID NO:13).

Figure 6 depicts the cDNA sequence (SEQ ID NO:14) and predicted amino acid sequence (SEQ ID NO:15) of human T125b, an alternatively spliced form of T125. The open reading frame of SEQ ID NO:14 extends from nucleotide 194 to nucleotide 934 of SEQ ID NO:14 (SEQ ID NO:16)

Figure 7 depicts the cDNA sequence (SEQ ID NO:17) and predicted amino acid sequence (SEQ ID NO:18) of T125c, an alternatively spliced form of human T125. The open reading frame of SEQ ID NO:17 extends from nucleotide 194 to nucleotide 823 of SEQ ID NO:17 (SEQ ID NO:19).

Detailed Description of the Invention

The present invention is based on the discovery of a cDNA molecule encoding human T125, a secreted protein.

A nucleotide sequence encoding a human T125 protein is shown in Figure 1 (SEQ ID NO:1; SEQ ID NO:3 includes the open reading frame only). A predicted amino acid sequence of T125 protein is also shown in Figure 1 (SEQ ID NO: 2). A cDNA sequence encoding a murine T125 protein is shown in Figure 4 (SEQ ID NO:8). The open reading frame only of this cDNA (nucleotides 13-837 of SEQ ID NO:8; SEQ ID NO:9) encodes a 275 amino acid protein (SEQ ID NO:10) that is also shown in Figure 4.

In addition, there are three additional alternatively spliced forms of human T125: T125a, T125b, and T125c. Figure 5 depicts the cDNA sequence (SEQ ID NO:11) and predicted amino acid sequence (SEQ ID NO:12) of human T125a. The open reading frame of SEQ ID NO:11 extends from nucleotide 94 to nucleotide 442 of SEQ ID NO:11 (SEQ ID NO:13). Figure 6 depicts the cDNA sequence (SEQ ID NO:14) and predicted amino acid sequence (SEQ ID

NO:15) of human T125b. The open reading frame of SEQ ID NO:14 extends from nucleotide 194 to nucleotide 934 of SEQ ID NO:14 (SEQ ID NO:16). Figure 7 depicts the cDNA sequence (SEQ ID NO:17) and predicted amino acid sequence (SEQ ID NO:18) of T125c. The open reading frame of SEQ ID NO:17 extends from nucleotide 194 to nucleotide 823 of SEQ ID NO:17 (SEQ ID NO:19).

Unless otherwise specified, "T125" (or "TANGO 125") is used to refer to all forms of T125 (T125, T125a, T125b, and T125c).

The T125 cDNA of Figure 1 (SEQ ID NO:1), which is approximately 1512 nucleotides long including untranslated regions, encodes a protein amino acid having a molecular weight of approximately 30 kDa (excluding post-translational modifications). A plasmid containing a cDNA encoding human T125 (with the plasmid name of pDH169) was deposited with American Type Culture Collection (ATCC), Rockville, Maryland on March 12, 1998 and assigned Accession Number 98693. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Sequence analysis revealed that T125 is homologous to GenBank entry gi-1841553, a protein having two EGF-like domains.

Alignment of the EGF-like domains of human T125 protein (SEQ ID NOS:5 and 6) with an EGF-like domain consensus sequence (SEQ ID NO:6) derived from a hidden Markov model revealed that the EGF1 domain is 44% identical (15/34 amino acids) and 65% similar (22/34 amino acids) and that the EGF2 domain is 35% identical (12/34 amino acids) and 71% similar (24/34 amino acids).

In general, EGF-like domains are found in the extracellular portion of membrane-bound proteins or in secreted proteins. EGF-like domains typically include six cysteine residues involved in disulfide bond formation with two conserved glycines between the fifth and sixth cysteine. The secondary structure of EGF-like domains appears to be a two-stranded ß-sheet followed by a loop to a C-terminal short two-stranded sheet.

T125 is expressed as a series of transcripts between 1.3 and 3 kb, which are expressed at various levels in the spleen, thymus, prostate, testes, ovary, small intestine, colon, heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas, the highest level of expression being observed in the placenta. T125 mRNA was not detected in peripheral blood leukocytes.

Human T125 is one member of a family of molecules (the "T125 family") having certain conserved structural and functional features. The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Such family members can be naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin and a homologue of that protein of murine origin, as well as a second, distinct protein of human origin and a murine homologue of that protein. Members of a family may also have common functional characteristics.

In one embodiment, a T125 protein includes an EGF-like domain having at least about 65%, preferably at least about 75%, and more preferably about 85%, 95%, or 98% amino acid sequence identity to the EGF-like domain of SEQ ID NOS:5 or 6.

Preferred T125 polypeptides of the present invention have an amino acid sequence sufficiently identical to the EGF-like domain amino acid sequence of SEQ ID NOS:5 or 6. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences have a common structural domain and/or common functional activity. For example, amino acid or nucleotide sequences which contain a common structural domain having about 65% identity, preferably 75% identity, more preferably 85%, 95%, or 98% identity are defined herein as sufficiently identical.

As used interchangeably herein a "T125 activity", "biological activity of T125" or "functional activity of T125", refers to an activity exerted by a T125 polypeptide or a nucleic acid molecule encoding a T125 polypeptide on a T125 responsive cell as determined *in vivo*, or *in vitro*, according to standard techniques. A T125 activity can be a direct activity, such as an association with or an enzymatic activity on a second protein or an indirect activity, such as a cellular signaling activity mediated by interaction of the T125 protein with a second protein (e.g., a T125 receptor).

Accordingly, another embodiment of the invention features isolated T125 proteins and polypeptides having a T125 activity.

Yet another embodiment of the invention features T125 molecules which contain a signal sequence. Generally, a signal sequence (or signal peptide) is a peptide containing about 20 amino acids which occurs at the extreme N-terminal end of secretory and integral

membrane proteins and which contains large numbers of hydrophobic amino acid residues and serves to direct a protein containing such a sequence to a lipid bilayer.

Various aspects of the invention are described in further detail in the following subsections.

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode T125 proteins or biologically active portions thereof, as well as nucleic acid molecules sufficient for use as hybridization probes to identify T125-encoding nucleic acids (e.g., T125 mRNA) and fragments for use as PCR primers for the amplification or mutation of T125 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences (preferably protein encoding sequences) which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated T125 nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide

sequence of SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the cDNA of ATCC 98693, or a complement of any of these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequences of SEQ ID NO:1, 3, 11, 13, 14, 16, 17, or 19, or the cDNA of ATCC 98693 as a hybridization probe, T125 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook et al., eds., *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A nucleic acid of the invention can be amplified using cDNA, mRNA or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to T125 nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693, or a portion thereof. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence encoding T125, for example, a fragment which can

be used as a probe or primer or a fragment encoding a biologically active portion of T125. The nucleotide sequence determined from the cloning of the human T125 gene allows for the generation of probes and primers designed for use in identifying and/or cloning T125 homologues in other cell types, e.g., from other tissues, as well as T125 homologues from other mammals. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 or 400 consecutive nucleotides of the sense or anti-sense sequence of SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693 or of a naturally occurring mutant of SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693.

Probes based on the human T125 nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or identical proteins. The probe comprises a label group attached thereto, e.g., a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which mis-express a T125 protein, such as by measuring a level of a T125-encoding nucleic acid in a sample of cells from a subject, e.g., detecting T125 mRNA levels or determining whether a genomic T125 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of T125" can be prepared by isolating a portion of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the cDNA of ATCC 98693 which encodes a polypeptide having a T125 biological activity, expressing the encoded portion of T125 protein (e.g., by recombinant

expression *in vitro*) and assessing the activity of the encoded portion of T125. For example, a nucleic acid fragment encoding a biologically active portion of T125 includes an EGF-like domain, e.g., SEQ ID NOS:5 or 6.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693 due to degeneracy of the genetic code and thus encode the same T125 protein as that encoded by the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693.

In addition to the human T125 nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of T125 may exist within a population (e.g., the human population). Such genetic polymorphism in the T125 gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a T125 protein, preferably a mammalian T125 protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the T125 gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in T125 that are the result of natural allelic variation and that do not alter the functional activity of T125 are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding T125 proteins from other species (T125 homologues), which have a nucleotide sequence which differs from that of a human T125, are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the T125 cDNA

of the invention can be isolated based on their identity to the human T125 nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. For example, a membrane-bound human T125 cDNA can be isolated based on its identity to soluble human T125.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 425 (450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1290) nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence, preferably the coding sequence, of SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, preferably 75%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, SEQ ID NO:3, the cDNA of ATCC 98693 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In addition to naturally-occurring allelic variants of the T125 sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, the cDNA of ATCC 98693, thereby leading to changes in the amino acid sequence of the encoded T125 protein, without altering the functional ability of the T125 protein. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of T125 (e.g., the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the T125 proteins of various species are predicted to be particularly unamenable to alteration.

For example, preferred T125 proteins of the present invention, contain at least one EGF-like domain. Such conserved domains are less likely to be amenable to mutation. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved among T125 of various species) may not be essential for activity and thus are likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding T125 proteins that contain changes in amino acid residues that are not essential for activity. Such T125 proteins differ in amino acid sequence from SEQ ID NO:2 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule includes a nucleotide sequence encoding a protein that includes an amino acid sequence that is at least about 45% identical, 65%, 75%, 85%, 95%,

or 98% identical to the amino acid sequence of SEQ ID NO:2.

An isolated nucleic acid molecule encoding a T125 protein having a sequence which differs from that of SEQ ID NO:2 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, the cDNA of ATCC 98693 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in T125 is preferably replaced with another amino acid residue from the same side chain family. Alternatively, mutations can be introduced randomly along all or part of a T125 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for T125 biological activity to identify mutants that

retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant T125 protein can be assayed for the ability to form protein:protein interactions. In yet another preferred embodiment, a mutant T125 can be assayed for the ability to modulate cellular proliferation or cellular differentiation.

The present invention encompasses antisense nucleic acid molecules, i.e., molecules which are complementary to a sense nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire T125 coding strand, or to only a portion thereof, e.g., all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can be antisense to a noncoding region of the coding strand of a nucleotide sequence encoding T125. The noncoding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

Given the coding strand sequences encoding T125 disclosed herein (e.g., SEQ ID NO:1 or SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of T125 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of T125 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of T125 mRNA, e.g., an oligonucleotide having

the sequence CTGAGAGCCCCTCATGGCCTGTGCCTCCAG (SEQ ID NO:8) or AGCCCCTCATGGCCTGTGCC (SEQ ID NO:9). An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil,

(acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a T125 protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the

control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave T125 mRNA transcripts to thereby inhibit translation of T125 mRNA. A ribozyme having specificity for a T125-encoding nucleic acid can be designed based upon the nucleotide sequence of a T125 cDNA disclosed herein (e.g., SEQ ID NO:1, SEQ ID NO:3). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a T125-encoding mRNA. See, e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742. Alternatively, T125 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel and Szostak (1993) *Science* 261:1411-1418.

The invention also encompasses nucleic acid molecules which form triple helical structures. For example, T125 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the T125 (e.g., the T125 promoter and/or enhancers) to form triple helical structures that prevent transcription of the T125 gene in target cells. See generally, Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

In preferred embodiments, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) *Bioorganic & Medicinal Chemistry* 4:5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

PNAs of T125 can be used therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of T125 can also be used, e.g., in the

analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as 'artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup (1996) *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup (1996) *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675).

In another embodiment, PNAs of T125 can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of T125 can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNAse H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) *supra* and Finn et al. (1996) *Nucleic Acids Research* 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag et al. (1989) *Nucleic Acid Res.* 17:5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) *Nucleic Acids Research* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment

(Peterser et al. (1975) *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol et al. (1988) *Bio/Techniques* 6:958-976) or intercalating agents (See, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

II. Isolated T125 Proteins and Anti-T125 Antibodies

One aspect of the invention pertains to isolated T125 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-T125 antibodies. In one embodiment, native T125 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, T125 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a T125 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from

the cell or tissue source from which the T125 protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of T125 protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, T125 protein that is substantially free of cellular material includes preparations of T125 protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of non-T125 protein (also referred to herein as a "contaminating protein"). When the T125 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When T125 protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, i.e., it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of T125 protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or non-T125 chemicals.

Biologically active portions of a T125 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the T125 protein (e.g., the amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4), which include less amino acids than the full length T125 proteins, and exhibit at least one activity of a T125 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the T125 protein. A biologically active portion of a T125 protein can be a polypeptide which is, for example,

10, 25, 50, 100 or more amino acids in length. Preferred biologically active polypeptides include one or more identified T125 structural domains, e.g., an EGF-like domain (SEQ ID NOS:5 or 6).

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native T125 protein. Preferred T125 protein has the amino acid sequence shown of SEQ ID NO:2. Other useful T125 proteins are substantially identical to SEQ ID NO:2 and retain the functional activity of the protein of SEQ ID NO:2 yet differ in amino acid sequence due to natural allelic variation or mutagenesis. Accordingly, a useful T125 protein is a protein which includes an amino acid sequence at least about 45%, preferably 55%, 65%, 75%, 85%, 95%, or 99% identical to the amino acid sequence of SEQ ID NO:2 and retains the functional activity of the T125 proteins of SEQ ID NO:2. In other instances, the T125 protein is a protein having an amino acid sequence 55%, 65%, 75%, 85%, 95%, or 98% identical to an EGF-like domain (SEQ ID NOS:5 or 6). In a preferred embodiment, the T125 protein retains the functional activity of the T125 protein of SEQ ID NO:2.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at

that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions x 100).

The determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to T125 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to T125 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The invention also provides T125 chimeric or fusion proteins. As used herein, a T125 "chimeric protein" or "fusion protein" comprises a T125 polypeptide operatively linked to a non-T125 polypeptide. A "T125 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to T125, whereas a "non-T125 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially identical to the T125 protein, e.g., a protein which is different from the T125 protein and which is derived from the same or a different organism. Within a T125 fusion protein the T125 polypeptide can correspond to all or a portion of a T125 protein, preferably at least one biologically active portion of a T125 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the T125 polypeptide and the non-T125 polypeptide are fused in-frame to each other. The non-T125 polypeptide can be fused to the N-terminus or C-terminus of the T125 polypeptide.

One useful fusion protein is a GST-T125 fusion protein in which the T125 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant T125.

In another embodiment, the fusion protein is a T125 protein containing a heterologous signal sequence at its N-terminus. For example, the native T125 signal sequence (i.e., about amino acids 1 to 22 of SEQ ID NO:2) can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of T125 can be

increased through use of a heterologous signal sequence. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (*Molecular cloning*, Sambrook et al, second edition, Cold spring harbor laboratory press, 1989) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

In yet another embodiment, the fusion protein is an T125-immunoglobulin fusion protein in which all or part of T125 is fused to sequences derived from a member of the immunoglobulin protein family (see, e.g., PCT Publication Number WO 88/07087 and Aruffo et al. (1990) *Cell* 61:1303-1313. The T125-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a T125 ligand and a protein with which it normally interacts. Moreover, the T125-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-T125 antibodies in a subject, to purify T125 ligands and in screening assays to identify molecules which inhibit the interaction of T125 with a T125 ligand.

Preferably, a T125 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini

for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, e.g., *Current Protocols in Molecular Biology*, Ausubel et al. eds., John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). An T125-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the T125 protein.

The present invention also pertains to variants of the T125 proteins which function as either T125 agonists (mimetics) or as T125 antagonists. Variants of the T125 protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the T125 protein. An agonist of the T125 protein can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the T125 protein. An antagonist of the T125 protein can inhibit one or more of the activities of the naturally occurring form of the T125 protein by, for example, competitively binding to a T125 receptor. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject

relative to treatment with the naturally occurring form of the T125 proteins.

Variants of the T125 protein which function as either T125 agonists (mimetics) or as T125 antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the T125 protein for T125 protein agonist or antagonist activity. In one embodiment, a variegated library of T125 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of T125 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential T125 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of T125 sequences therein. There are a variety of methods which can be used to produce libraries of potential T125 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential T125 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of the T125 protein coding sequence can be used to generate a variegated population of T125 fragments for screening and subsequent selection of variants of a T125 protein. In

one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a T125 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the T125 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of T125 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify T125 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(3):327-331).

An isolated T125 protein, or a portion or fragment thereof, can be used as an immunogen to generate

antibodies that bind T125 using standard techniques for polyclonal and monoclonal antibody preparation. The full-length T125 protein can be used or, alternatively, the invention provides antigenic peptide fragments of T125 for use as immunogens. The antigenic peptide of T125 comprises at least 8 (preferably 10, 15, 20, or 30) amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an epitope of T125 such that an antibody raised against the peptide forms a specific immune complex with T125.

Preferred epitopes encompassed by the antigenic peptide are regions of T125 that are located on the surface of the protein, e.g., hydrophilic regions. A hydrophobicity analysis of the human T125 protein sequence indicates that the regions between, e.g., amino acids 85 and 95, between amino acids 210 and 220, and between amino acids 235 and 245 of SEQ ID NO:2 are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production.

A T125 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed T125 protein or a chemically synthesized T125 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic T125 preparation induces a polyclonal anti-T125 antibody response.

Accordingly, another aspect of the invention pertains to anti-T125 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin

molecules, i.e., molecules that contain an antigen binding site which specifically binds an antigen, such as T125. A molecule which specifically binds to T125 is a molecule which binds T125, but does not substantially bind other molecules in a sample, e.g., a biological sample, which naturally contains T125. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind T125. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of T125. A monoclonal antibody composition thus typically displays a single binding affinity for a particular T125 protein with which it immunoreacts.

Polyclonal anti-T125 antibodies can be prepared as described above by immunizing a suitable subject with a T125 immunogen. The anti-T125 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized T125. If desired, the antibody molecules directed against T125 can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the anti-T125 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (Kozbor et al. (1983) *Immunol*

Today 4:72), the EBV-hybridoma technique (Cole et al. (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing various antibodies monoclonal antibody hybridomas is well known (see, e.g., *Current Protocols in Immunology* (1994) Coligan et al. (eds.) John Wiley & Sons, Inc., New York, NY). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a T125 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds T125.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-T125 monoclonal antibody (see, e.g., Current Protocols in Immunology, *supra*; Galfre et al. (1977) *Nature* 266:55052; R.H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); and Lerner (1981) *Yale J. Biol. Med.*, 54:387-402. Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line, e.g., a myeloma cell line that is sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines.

These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind T125, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-T125 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with T125 to thereby isolate immunoglobulin library members that bind T125. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246:1275-1281; Griffiths et al. (1993) *EMBO J* 12:725-734.

Additionally, recombinant anti-T125 antibodies, such as chimeric and humanized monoclonal antibodies,

comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better et al. (1988) *Science* 240:1041-1043; Liu et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu et al. (1987) *J. Immunol.* 139:3521-3526; Sun et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura et al. (1987) *Canc. Res.* 47:999-1005; Wood et al. (1985) *Nature* 314:446-449; and Shaw et al. (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison, (1985) *Science* 229:1202-1207; Oi et al. (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones et al. (1986) *Nature* 321:552-525; Verhoeyan et al. (1988) *Science* 239:1534; and Beidler et al. (1988) *J. Immunol.* 141:4053-4060.

An anti-T125 antibody (e.g., monoclonal antibody) can be used to isolate T125 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-T125 antibody can facilitate the purification of natural T125 from cells and of recombinantly produced T125 expressed in host cells. Moreover, an anti-T125 antibody can be used to detect T125 protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the T125 protein. Anti-T125 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable

substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding T125 (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are

replicated along with the host genome. Moreover, certain vectors, expression vectors, are capable of directing the expression of genes to which they are operatively linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be

appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., T125 proteins, mutant forms of T125, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of T125 in prokaryotic or eukaryotic cells, e.g., bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the

fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration

of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the T125 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari et al. (1987) *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al. (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, T125 can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook et al. (*supra*).

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of

suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to T125 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the

control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes See Weintraub et al., *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, T125 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding T125 or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) T125 protein. Accordingly, the invention further provides methods for producing T125 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding T125 has been introduced) in a suitable medium such that T125 protein is produced. In another embodiment, the method further comprises isolating T125 from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which T125-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic

animals in which exogenous T125 sequences have been introduced into their genome or homologous recombinant animals in which endogenous T125 sequences have been altered. Such animals are useful for studying the function and/or activity of T125 and for identifying and/or evaluating modulators of T125 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous T125 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing T125-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The T125 cDNA sequence e.g., that of (SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 98693) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human T125 gene, such as a mouse T125 gene, can be isolated based on hybridization to the human T125 cDNA and used as

a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the T125 transgene to direct expression of T125 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the T125 transgene in its genome and/or expression of T125 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding T125 can further be bred to other transgenic animals carrying other transgenes.

To create an homologous recombinant animal, a vector is prepared which contains at least a portion of a T125 gene (e.g., a human or a non-human homolog of the T125 gene, e.g., a murine T125 gene) into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the T125 gene. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous T125 gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous T125 gene is mutated or otherwise altered but still encodes

functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous T125 protein). In the homologous recombination vector, the altered portion of the T125 gene is flanked at its 5' and 3' ends by additional nucleic acid of the T125 gene to allow for homologous recombination to occur between the exogenous T125 gene carried by the vector and an endogenous T125 gene in an embryonic stem cell. The additional flanking T125 nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas and Capecchi (1987) *Cell* 51:503 for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced T125 gene has homologously recombined with the endogenous T125 gene are selected (see e.g., Li et al. (1992) *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see, e.g., Bradley in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 and in

PCT Publication Nos. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut et al. (1997) *Nature* 385:810-813 and PCT Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_o phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the

animal from which the cell, e.g., the somatic cell, is isolated.

IV. Pharmaceutical Compositions

The T125 nucleic acid molecules, T125 proteins, and anti-T125 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or

sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such

as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a T125 protein or anti-T125 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an

excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring. For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods

for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see e.g., Chen et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector

can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) detection assays (e.g., chromosomal mapping, tissue typing, forensic biology), c) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and d) methods of treatment (e.g., therapeutic and prophylactic). A T125 protein interacts with other cellular proteins and can thus be used for (i) regulation of cellular proliferation; (ii) regulation of cellular differentiation; and (iii) regulation of cell survival. The isolated nucleic acid molecules of the invention can be used to express T125 protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect T125 mRNA (e.g., in a biological sample) or a genetic lesion in a T125 gene, and to modulate T125 activity. In addition, the T125 proteins can be used to screen drugs or compounds which modulate the T125 activity or expression as well as to treat disorders characterized by insufficient or excessive production of T125 protein or production of T125 protein forms which have decreased or aberrant activity compared to T125 wild type protein. In addition, the anti-T125 antibodies of the invention can be used to detect and isolate T125 proteins and modulate T125 activity.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

A. Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to T125 proteins or have a stimulatory or inhibitory effect on, for example, T125 expression or T125 activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a T125 polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al. (1994). *J. Med. Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et

al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop et al. (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Bio/Techniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (U.S. Patent No. 5,223,409), spores (Patent Nos. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382; and Felici (1991) *J. Mol. Biol.* 222:301-310).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a T125 protein, or a biologically active portion thereof, is contacted with a test compound and the ability of the test compound to bind to a T125 protein determined. The cell, for example, can be a yeast cell or a cell of mammalian origin. Determining the ability of the test compound to bind to the T125 protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the T125 protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In a preferred embodiment, the assay comprises contacting a cell which

expresses a T125 protein, or a biologically active portion thereof, with a known compound which binds T125 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a T125 protein, wherein determining the ability of the test compound to interact with a T125 protein comprises determining the ability of the test compound to preferentially bind to T125 or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a T125 protein, or a biologically active portion thereof, with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the T125 protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of T125 or a biologically active portion thereof can be accomplished, for example, by determining the ability of the T125 protein to bind to or interact with a T125 target molecule. As used herein, a "target molecule" is a molecule with which a T125 protein binds or interacts in nature, for example, a molecule associated with the outer surface of a cell membrane or another secreted molecule. A T125 target molecule can be a non-T125 molecule or a T125 protein or polypeptide of the present invention.

Determining the ability of the T125 protein to bind to or interact with a T125 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the T125 protein to bind to or interact with a T125 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be

determined by detecting induction of a cellular second messenger of the target (e.g., intracellular Ca^{2+}, diacylglycerol, IP3, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (e.g., a T125-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g. luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a T125 protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the T125 protein or biologically active portion thereof. Binding of the test compound to the T125 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the T125 protein or biologically active portion thereof with a known compound which binds T125 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a T125 protein, wherein determining the ability of the test compound to interact with a T125 protein comprises determining the ability of the test compound to preferentially bind to T125 or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting T125 protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the T125 protein or biologically active portion thereof. Determining the ability of the test compound to modulate

the activity of T125 can be accomplished, for example, by determining the ability of the T125 protein to bind to a T125 target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of T125 can be accomplished by determining the ability of the T125 protein further modulate a T125 target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In some cell-free assays, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of T125 is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either T125 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to T125, or interaction of T125 with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a

domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/T125 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical; St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or T125 protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of T125 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either T125 or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated T125 or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with T125 or target molecules but which do not interfere with binding of the T125 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or T125 trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of

complexes using antibodies reactive with the T125 or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the T125 or target molecule.

In another embodiment, modulators of T125 expression are identified in a method in which a cell is contacted with a candidate compound and the expression of T125 mRNA or protein in the cell is determined. The level of expression of T125 mRNA or protein in the presence of the candidate compound is compared to the level of expression of T125 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of T125 expression based on this comparison. For example, when expression of T125 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of T125 mRNA or protein expression. Alternatively, when expression of T125 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of T125 mRNA or protein expression. The level of T125 mRNA or protein expression in the cells can be determined by methods described herein for detecting T125 mRNA or protein.

In yet another aspect of the invention, the T125 proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Bio/Techniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and WO 94/10300), to identify other proteins, which bind to or interact with

T125 ("T125-binding proteins" or "T125-bp") and modulate T125 activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for T125 is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an T125-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with T125.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic

the activity of T125 can be accomplished, for example, by determining the ability of the T125 protein to bind to a T125 target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of T125 can be accomplished by determining the ability of the T125 protein further modulate a T125 target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In some cell-free assays, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of T125 is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either T125 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to T125, or interaction of T125 with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a

domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/T125 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical; St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or T125 protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of T125 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either T125 or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated T125 or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with T125 or target molecules but which do not interfere with binding of the T125 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or T125 trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of

complexes using antibodies reactive with the T125 or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the T125 or target molecule.

In another embodiment, modulators of T125 expression are identified in a method in which a cell is contacted with a candidate compound and the expression of T125 mRNA or protein in the cell is determined. The level of expression of T125 mRNA or protein in the presence of the candidate compound is compared to the level of expression of T125 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of T125 expression based on this comparison. For example, when expression of T125 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of T125 mRNA or protein expression. Alternatively, when expression of T125 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of T125 mRNA or protein expression. The level of T125 mRNA or protein expression in the cells can be determined by methods described herein for detecting T125 mRNA or protein.

In yet another aspect of the invention, the T125 proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Bio/Techniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and WO 94/10300), to identify other proteins, which bind to or interact with

T125 ("T125-binding proteins" or "T125-bp") and modulate T125 activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for T125 is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an T125-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with T125.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic

disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. Accordingly, T125 nucleic acid molecules described herein or fragments thereof, can be used to map the location of T125 genes on a chromosome. The mapping of the T125 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, T125 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the T125 sequences. Computer analysis of T125 sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the T125 sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line

in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio et al. (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the T125 sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a T125 sequence to its chromosome include *in situ* hybridization (described in Fan et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection.

Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma et al., Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, e.g., Egeland et al. (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the T125 gene can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes such as deletions or translocations that are visible from

chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The T125 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the T125 sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from

individuals and from tissue. The T125 sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from T125 sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of Partial T125 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used

to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the T125 sequences or portions thereof, e.g., fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 or 30 bases.

The T125 sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such T125 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, e.g., T125 primers or probes can be used to screen tissue culture

for contamination (i.e., screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trails are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining T125 protein and/or nucleic acid expression as well as T125 activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant T125 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with T125 protein, nucleic acid expression or activity. For example, mutations in a T125 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with T125 protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining T125 protein, nucleic acid expression or T125 activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to

determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs or other compounds) on the expression or activity of T125 in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of T125 in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting T125 protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes T125 protein such that the presence of T125 is detected in the biological sample. A preferred agent for detecting T125 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to T125 mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length T125 nucleic acid, such as the nucleic acid of SEQ ID NO: 1 or 3, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to T125 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting T125 protein is an antibody capable of binding to T125 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the

probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect T125 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of T125 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of T125 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of T125 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of T125 protein include introducing into a subject a labeled anti-T125 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting T125 protein, mRNA, or genomic DNA, such that the presence of T125 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of T125 protein, mRNA or genomic DNA in the control sample with the presence of T125 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of T125 in a biological sample (a test sample). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing a disorder associated with aberrant expression of T125. For example, the kit can comprise a labeled compound or agent capable of detecting T125 protein or mRNA in a biological sample and means for determining the amount of T125 in the sample (e.g., an anti-T125 antibody or an oligonucleotide probe which binds to DNA encoding T125, e.g., SEQ ID NO:1 or SEQ ID NO:3). Kits may also include instruction for observing that the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of T125 if the amount of T125 protein or mRNA is above or below a normal level.

For antibody-based kits, the kit may comprise, for example: (1) a first antibody (e.g., attached to a solid support) which binds to T125 protein; and, optionally, (2) a second, different antibody which binds to T125 protein or the first antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit may comprise, for example: (1) a oligonucleotide, e.g., a detectably labelled oligonucleotide, which hybridizes to a T125 nucleic acid sequence or (2) a pair of primers useful for amplifying a T125 nucleic acid molecule.

The kit may also comprise, e.g., a buffering agent, a preservative, or a protein stabilizing agent. The kit may also comprise components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit may also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for observing whether the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression or activity of T125.

2. Prognostic Assays

The methods described herein can furthermore be utilized as diagnostic or prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with aberrant T125 expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with T125 protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing such a disease or disorder. Thus, the present invention provides a method in which a test sample is obtained from a subject and T125 protein or nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of T125 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant T125 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant T125 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with a specific agent or class of agents (e.g., agents of a type which decrease T125 activity). Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a

disorder associated with aberrant T125 expression or activity in which a test sample is obtained and T125 protein or nucleic acid is detected (e.g., wherein the presence of T125 protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant T125 expression or activity).

The methods of the invention can also be used to detect genetic lesions or mutations in a T125 gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a T125-protein, or the mis-expression of the T125 gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a T125 gene; 2) an addition of one or more nucleotides to a T125 gene; 3) a substitution of one or more nucleotides of a T125 gene, 4) a chromosomal rearrangement of a T125 gene; 5) an alteration in the level of a messenger RNA transcript of a T125 gene, 6) aberrant modification of a T125 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a T125 gene, 8) a non-wild type level of a T125-protein, 9) allelic loss of a T125 gene, and 10) inappropriate post-translational modification of a T125-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a T125 gene. A preferred biological sample is a peripheral

blood leukocyte sample isolated by conventional means from a subject.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science* 241:1077-1080; and Nakazawa et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the T125-gene (see Abravaya et al. (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a T125 gene under conditions such that hybridization and amplification of the T125-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al. (1988) *Bio/Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection

schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a T125 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, e.g., U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in T125 can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin et al. (1996) *Human Mutation* 7:244-255; Kozal et al. (1996) *Nature Medicine* 2:753-759). For example, genetic mutations in T125 can be identified in two-dimensional arrays containing light-generated DNA probes as described in Cronin et al. supra. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to

the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the T125 gene and detect mutations by comparing the sequence of the sample T125 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Bio/Techniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT Publication No. WO 94/16101; Cohen et al. (1996) *Adv. Chromatogr.* 36:127-162; and Griffin et al. (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the T125 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al. (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type T125 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and

with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, e.g., Cotton et al (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba et al (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in T125 cDNAs obtained from samples of cells. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a T125 sequence, e.g., a wild-type T125 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, e.g., U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in T125 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) *Proc Natl. Acad. Sci USA*: 86:2766; see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control T125 nucleic acids will be denatured and allowed to renature. The secondary structure of

single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) *Trends Genet* 7:5).

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al. (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) *Nature* 324:163); Saiki et al. (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of

different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al. (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition, it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al. (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a T125 gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which T125 is expressed

may be utilized in the prognostic assays described herein.

3. Pharmacogenomics

Agents, or modulators which have a stimulatory or inhibitory effect on T125 activity (e.g., T125 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (e.g., an immunological disorder) associated with aberrant T125 activity. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of T125 protein, expression of T125 nucleic acid, or mutation content of T125 genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Linder (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions

transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular

basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of T125 protein, expression of T125 nucleic acid, or mutation content of T125 genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a T125 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

4. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of T125 (e.g., the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase T125 gene expression, protein levels, or upregulate T125 activity, can be monitored in clinical trails of subjects exhibiting decreased T125 gene expression, protein levels, or downregulated T125 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease T125 gene expression, protein levels, or downregulated T125 activity, can be monitored in clinical trails of subjects exhibiting increased T125 gene expression, protein levels, or upregulated T125 activity. In such clinical trials, the expression or

activity of T125 and, preferably, other genes that have been implicated in, for example, a cellular proliferation disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

For example, and not by way of limitation, genes, including T125, that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) which modulates T125 activity (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of T125 and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of T125 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a T125 protein, mRNA, or genomic DNA in the preadministration sample;

(iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the T125 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the T125 protein, mRNA, or genomic DNA in the pre-administration sample with the T125 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of T125 to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of T125 to lower levels than detected, i.e., to decrease the effectiveness of the agent.

C. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant T125 expression or activity. Such disorders include wound healing and cancer.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant T125 expression or activity, by administering to the subject an agent which modulates T125 expression or at least one T125 activity. Subjects at risk for a disease which is caused or contributed to by aberrant T125 expression or activity can be identified by, for example, any or a combination of diagnostic or

prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the T125 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of T125 aberrancy, for example, a T125 agonist or T125 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating T125 expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of T125 protein activity associated with the cell. An agent that modulates T125 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a T125 protein, a peptide, a T125 peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more of the biological activities of T125 protein. Examples of such stimulatory agents include active T125 protein and a nucleic acid molecule encoding T125 that has been introduced into the cell. In another embodiment, the agent inhibits one or more of the biological activities of T125 protein. Examples of such inhibitory agents include antisense T125 nucleic acid molecules and anti-T125 antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant

expression or activity of a T125 protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) T125 expression or activity. In another embodiment, the method involves administering a T125 protein or nucleic acid molecule as therapy to compensate for reduced or aberrant T125 expression or activity.

Stimulation of T125 activity is desirable in situations in which T125 is abnormally downregulated and/or in which increased T125 activity is likely to have a beneficial effect. Conversely, inhibition of T125 activity is desirable in situations in which T125 is abnormally upregulated and/or in which decreased T125 activity is likely to have a beneficial effect.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

EXAMPLES

Example 1: Isolation and Characterization of Human T125 cDNAs

Human aortic endothelial cells (obtained from Clonetics Corporation; San Diego, CA) were expanded in culture with Endothelial Cell Growth Media (EGM; Clonetics) according to the recommendations of the supplier. When the cells reached ~80-90% confluence, they were stimulated with TNF (10 ng/ml) and cycloheximide (CHI; 40 micrograms/ml) for 4 hours. Total RNA was isolated using the RNeasy Midi Kit (Qiagen;

Chatsworth, CA), and the poly A+ fraction was further purified using Oligotex beads (Qiagen).

Three micrograms of poly A+ RNA were used to synthesize a cDNA library using the Superscript cDNA Synthesis kit (Gibco BRL; Gaithersburg, MD). Complementary DNA was directionally cloned into the expression plasmid pMET7 using the SalI and NotI sites in the polylinker to construct a plasmid library. Transformants were picked and grown up for single-pass sequencing. A partial cDNA clone (jthdc042c10) was identified that encoded a protein with homology to a Genbank entry (gi-1841553) which appeared to encode a secreted protein with two EGF domains (note: This Genbank entry seems to be a condensation of genomic sequence relying on EST sequence to define the coding region, and there may be some errors or alternative splicing within the entry.) Jthdc042c10 was completely sequenced, and lacked an appropriate start codon. Therefore additional homologous clones in the library were identified by database searches and sequenced. One clone (jthdc054a01) contained a 273 amino acid open reading frame that was ~37% identical with gi-1841553, and contained a predicted signal sequence (amino acids 1-22). Two regions of Tango 125 showed similarity to EGF domains (amino acids 107-134 and amino acids 141-176 of SEQ ID NO:2), and there was complete conservation of all cysteines between Tango 125 and gi-1841553.

Example 2: <u>Distribution of T125 mRNA in Human Tissues</u>

The expression of T125 was analyzed using Northern blot hybridization. Primers (5' GCTCACGGGGACCCTGTC 3' (SEQ ID NO:10) and 5' CAGTGCCTGCGAGGCCAG 3' (SEQ ID NO:11)) were used to amplify a 585 bp fragment from the 5' end of the T125 coding region. The DNA was radioactively labeled with ^{32}P-dCTP using the Prime-It kit

(Stratagene, La Jolla, CA) according to the instructions of the supplier. Filters containing human mRNA (MTNI and MTNII from Clontech, Palo Alto, CA) were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

T125 is expressed as series of transcripts between 1.3 and 3 kb. These transcripts are found at variable levels in all tissues examined (spleen, thymus, prostate, testes, ovary, small intestine, colon, heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas) with the exception of peripheral blood leukocytes in which expression was not detected. The highest levels of T125 expression were observed in the placenta as a 3 kb transcript, with the next highest levels found in spleen and testis as ~2 and 1.5 kb transcripts respectively.

The various size transcripts seen on the Northern blots could be consistent with alternative splicing of the T125 gene. Although there were no changes in the coding region between the clones that were sequenced, the clones appeared to be partially spliced transcripts. It is unknown at this point if the alternative splicing is important for the regulation of expression, or whether additional clones containing variations in the coding sequence may also be expressed.

Human *in situ* expression analysis revealed that T125 is expressed in lung (ubiquitous with multifocal areas of higher expression), thymus (ubiquitous with multifocal areas of higher expression), heart, kidney, liver, non-follicular regions of the spleen. Expression was also observed, at a lower level, in brain and placenta. *In situ* expression analysis of human embryonic tissues revealed that T125 is expressed in most tissues

with the highest expression in heart, lung, kidney, and early fetal liver (E13.5 through E15.5)

Example 3: Characterization of T125 Proteins

In this example, the predicted amino acid sequence of human T125 protein was compared to amino acid sequences of known proteins and various motifs were identified. In addition, the molecular weight of the human T125 proteins was predicted.

The human T125 cDNA isolated as described above (Figure 1; SEQ ID NO:1) encodes a 273 amino acid protein (Figure 1; SEQ ID NO:2). The signal peptide prediction program SIGNALP (Nielsen et al. (1997) *Protein Engineering* 10:1-6) predicted that T125 includes a 22 amino acid signal peptide (amino acid 1 to about amino acid 22 of SEQ ID NO:2) preceding the 252 amino acid mature protein (about amino acid 23 to amino acid 274; SEQ ID NO:4).

As shown in Figure 2, T125 has two regions of homology (amino acids 107 to 134 of SEQ ID NO:2; SEQ ID NO:5; and amino acids 141 to 176 of SEQ ID NO:2; SEQ ID NO:6) to the EGF-like domain consensus sequence (SEQ ID NO:7). Both regions contain the six conserved cysteines and two conserved glycines between the fifth and sixth cysteine in the consensus sequence. The mature T125 protein is predicted to have a MW of 30 kDa (27 kDa without the signal peptide).

Example 4: Alternatively Spliced Forms of Human T125

Additional analysis revealed that the human T125 cDNA shown in Figure 1 represents one of four alternatively spliced forms of human T125. The three additional forms, T125a, T125b, and T125c are depicted in Figure 5, Figure 6, and Figure 7 respectively. Figure 5 depicts the cDNA sequence (SEQ ID NO:11) and predicted

amino acid sequence (SEQ ID NO:12) of human T125a. The open reading frame of SEQ ID NO:11 extends from nucleotide 94 to nucleotide 442 of SEQ ID NO:11 (SEQ ID NO:13). Figure 6 depicts the cDNA sequence (SEQ ID NO:14) and predicted amino acid sequence (SEQ ID NO:15) of human T125b. The open reading frame of SEQ ID NO:14 extends from nucleotide 194 to nucleotide 934 of SEQ ID NO:14 (SEQ ID NO:16). Figure 7 depicts the cDNA sequence (SEQ ID NO:17) and predicted amino acid sequence (SEQ ID NO:18) of T125c. The open reading frame of SEQ ID NO:17 extends from nucleotide 194 to nucleotide 823 of SEQ ID NO:17 (SEQ ID NO:19).

The four forms arise from the use of three exons. All four forms include exon 1. The form of human T125 (called T125) depicted in Figure 1 includes exon 2 and exon 3 in addition to exon 1. T125a includes exon 2 in addition to exon 1. T125b includes exon 1 only. T125c includes exon 3 in addition to exon 1. The coding sequence of both T125b and T125c begins at an ATG that is upstream of the ATG that is the beginning of the coding sequence for T125 and T125b. T125 may be subject to two types of post-transcriptional regulation: choice of initiation site and choice of splicing.

Example 5: Identification of Murine T125 and Distribution of T125 in Murine Tissue

A full-length murine T125 cDNA clone was isolated. This 846 nucleotide cDNA is depicted in Figure 4 (SEQ ID NO:8). The open reading frame of this molecule extends from nucleotide 13 to nucleotide 837 of SEQ ID NO:8 (SEQ ID NO:9) and encodes a 275 amino acid protein (SEQ ID NO:10).

Northern blot analysis revealed that murine T125 is expressed at a moderate level in heart, lung, and liver and at a lower level in brain and kidney.

In situ expression analysis revealed that murine T125 is expressed in lung (ubiquitous with multifocal areas of higher expression), thymus (ubiquitous with multifocal areas of higher expression), liver (ubiquitous with probable expression in hepatocytes), kidney (ubiquitous), spleen (non-follicular), brain (low, but ubiquitous), placenta (ubiquitous, inner mass). *In situ* expression analysis of murine embryonic tissue revealed ubiquitous expression at E13.5 through E15.5, with higher expression in lung, heart, liver, and kidney. At E16.5 through E18.5 and at P1.5, the ubiquitous expression of T125 decreases with higher signal persisting in lung, heart, and kidney.

Overexpression of murine T125 in mice using a retroviral expression system revealed the T125 overexpression may reduce triglyceride levels by nearly 50%.

Example 6: Preparation of T125 Proteins

Recombinant T125 can be produced in a variety of expression systems. For example, the mature T125 polypeptide can be expressed as a recombinant glutathione-S-transferase (GST) fusion protein in *E. coli*, and the fusion protein can be isolated and characterized. Specifically, as described above, T125 can be fused to GST, and this fusion protein can be expressed in *E. coli* strain PEB199. Expression of the GST-T125 fusion protein in PEB199 can be induced with IPTG. The recombinant fusion protein can be purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography using glutathione beads.

Example 7: Creation of Flag-tagged T125

A flag epitope-tagged version of T125 is constructed by PCR amplifying a T125 gene using a 3' primer that includes a nucleotide sequence encoding the DYKDDDDK flag epitope (SEQ ID NO:22) followed by a termination codon. The amplified clone is inserted into a pMET vector and the resulting construct is used to transiently transfected into HEK 293T cells in 150mM plates using Lipofectamine (GIBCO/BRL, Gaithersburg MD) according to the manufacturer's protocol. The cells are used to express flag-tagged T!@%.

Example 8: Retroviral Delivery of T125

Full length human or murine T125 is expressed *in vivo* mediated by retroviral infection. A sequence encoding a selected T125 is cloned into the retroviral vector MSCVneo (Hawley et al. (1994) *Gene Therapy* 1:136-138), and sequence verified. Bone marrow from 5-fluorouracil treated mice infected with the retrovirus is then transplanted into irradiated mouse recipients.

Example 9: T125 alkaline phosphatase N-terminal fusion protein

A vector expression a T125-alkaline phosphatase fusion protein is prepared by ligating a sequence encoding a selected T125 into AP-Tag3 vector (Tartaglia et al. (1995) *Cell* 83:1263-1271). The full-length open-reading frame of T125 is PCR amplified using a 5' primer incorporating a BglII restriction site prior to the nucleotides encoding the first amino acids of T125 and a 3' primer including a XhoI restriction site immediately following the termination codon of T125. Thus the open reading frame of the complete construct includes the complete sequence of human placental

alkaline phosphatase, including the signal peptide, followed by T125 sequence.

The resulting vector is transiently transfected into HEK 293T cells in 150mM plates using Lipofectamine (GIBCO/BRL) according to the manufacturer's protocol. Seventy-two hours post-transfection, the serum-free conditioned media (OptiMEM, GIBCO/BRL) is harvested, spun and filtered. Alkaline phosphatase activity in conditioned media is quantitated using an enzymatic assay kit (Phospha-Light, Tropix Inc.) according to the manufacturer's instructions. Conditioned medium samples are analyzed by SDS-PAGE followed by Western blot using anti-human alkaline phosphatase antibodies diluted 1:250 (Genzyme Corp., Cambridge MA) and detected by chemiluminescence.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed is:

1. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, or a complement thereof;

b) a nucleic acid molecule comprising a fragment of at least 500 nucleotides of the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, or a complement thereof;

c) nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693;

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or SEQ ID NO:4 or the polypeptide encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693; and

e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, which is selected from the group consisting of:

a) a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, or a complement thereof; and

b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693.

3. The nucleic acid molecule of claim 1 further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of claim 1.

6. The host cell of claim 4 which is a mammalian host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:

a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or SEQ ID NO:4;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by

the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions;

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3.

9. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693.

10. The polypeptide of claim 8 further comprising heterologous amino acid sequences.

11. An antibody which selectively binds to a polypeptide of claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693;

b) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or SEQ ID NO:4 or an amino

acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693; and

c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions;

comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

13. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98693.

14. A method for detecting the presence of a polypeptide of claim 8 in a sample, comprising:

a) contacting the sample with a compound which selectively binds to a polypeptide of claim 8; and

b) determining whether the compound binds to the polypeptide in the sample.

15. The method of claim 14, wherein the compound which binds to the polypeptide is an antibody.

16. A kit comprising a compound which selectively binds to a polypeptide of claim 8 and instructions for use.

17. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

18. The method of claim 17, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

19. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

20. A method for identifying a compound which binds to a polypeptide of claim 8 comprising the steps of:

a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and

b) determining whether the polypeptide binds to the test compound.

21. The method of claim 20, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detecting of test compound/polypeptide binding;

b) detection of binding using a competition binding assay;

c) detection of binding using an assay for EGF activity.

22. A method for modulating the activity of a polypeptide of claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim 8 with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

23. A method for identifying a compound which modulates the activity of a polypeptide of claim 8, comprising:

a) contacting a polypeptide of claim 8 with a test compound; and

b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

```
GTCGACCCACGCGTCCGCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGCTAGGGTCCATCT   79
CCAGTCCCAGGACACAGCAGGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGG  158
CACAGGTGGCCCCCCACCACCCGGAGGAGCAGCTCCTGTCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACC  237
CAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCC ATG AGG GGC TCT CAG GAG GTG CTG CTG ATG
                                      Met Arg Gly Ser Gln Glu Val Leu Leu Met
                                        1               5                  10

TGG CTT CTG GTG TTG GCA GTG GGC GGC ACA GAG CAC GCC TAC CGG CCC GGC CGT AGG GTG  363
Trp Leu Leu Val Leu Ala Val Gly Gly Thr Glu His Ala Tyr Arg Pro Gly Arg Arg Val
            15                  20                  25                  30

TGT GCT GTC CGG GCT CAC GGG GAC CCT GTC TCC GAG TCG TTC GTG CAG CGT GTG TAC CAG  423
Cys Ala Val Arg Ala His Gly Asp Pro Val Ser Glu Ser Phe Val Gln Arg Val Tyr Gln
        35                  40                  45                  50

CCC TTC CTC ACC ACC TGC GAC GGG CAC CGG GCC TGC AGC ACC TAC CGA ACC ATC TAT AGG  483
Pro Phe Leu Thr Thr Cys Asp Gly His Arg Ala Cys Ser Thr Tyr Arg Thr Ile Tyr Arg
    55                  60                  65                  70

ACC GCC TAC CGC CGC AGC CCT GGG CTG GCC CCT GCC AGG CCT CGC TAC GCG TGC TGC CCC  543
Thr Ala Tyr Arg Arg Ser Pro Gly Leu Ala Pro Ala Arg Pro Arg Tyr Ala Cys Cys Pro
                75                  80                  85                  90

GGC TGG AAG AGG ACC AGC GGG CTT CCT GGG GCC TGT GGA GCA GCA ATA TGC CAG CCG CCA  603
Gly Trp Lys Arg Thr Ser Gly Leu Pro Gly Ala Cys Gly Ala Ala Ile Cys Gln Pro Pro
            95                  100                 105                 110

TGC CGG AAC GGA GGG AGC TGT GTC CAG CCT GGC CGC TGC CGC TGC CCT GCA GGA TGG CGG  663
Cys Arg Asn Gly Gly Ser Cys Val Gln Pro Gly Arg Cys Arg Cys Pro Ala Gly Trp Arg
        115                 120                 125                 130

GGT GAC ACT TGC CAG TCA GAT GTG GAT GAA TGC AGT GCT AGG AGG GGC GGC TGT CCC CAG  723
Gly Asp Thr Cys Gln Ser Asp Val Asp Glu Cys Ser Ala Arg Arg Gly Gly Cys Pro Gln
    135                 140                 145                 150

CGC TGC ATC AAC ACC GCC GGC AGT TAC TGG TGC CAG TGT TGG GAG GGG CAC AGC CTG TCT  783
Arg Cys Ile Asn Thr Ala Gly Ser Tyr Trp Cys Gln Cys Trp Glu Gly His Ser Leu Ser
                155                 160                 165                 170
```

FIG. 1A

```
GCA GAC GGT ACA CTC TGT GTG CCC AAG GGA GGG CCC CCC AGG GTG GCC CCC AAC CCG ACA   843
Ala Asp Gly Thr Leu Cys Val Pro Lys Gly Gly Pro Pro Arg Val Ala Pro Asn Pro Thr
        175                 180                 185                 190

GGA GTG GAC AGT GCA ATG AAG GAA GTG CAG AGG CTG CAG TCC AGG GTG GAC CTG CTG   903
Gly Val Asp Ser Ala Met Lys Glu Val Gln Arg Leu Gln Ser Arg Val Asp Leu Leu
                195                 200                 205                 210

GAG GAG AAG CTG CAG CTG GTG CCC CCA CTG CAC AGC CTG GCC TCG CAG GCA CTG GAG   963
Glu Glu Lys Leu Gln Leu Val Leu Ala Pro Leu His Ser Leu Ala Ser Gln Ala Leu Glu
                215                 220                 225                 230

CAT GGG CTC CCG GAC CCC GGC AGC CTC CTG GTG CAC TCC TTC CAG CAG CTC GGC CGC ATC  1023
His Gly Leu Pro Asp Pro Gly Ser Leu Leu Val His Ser Phe Gln Gln Leu Gly Arg Ile
                235                 240                 245                 250

GAC TCC CTG AGC GAG CAG ATT TCC TTC CTG GAG GAG CAG CTG GGG TCC TGC TCC TGC AAG  1083
Asp Ser Leu Ser Glu Gln Ile Ser Phe Leu Glu Glu Gln Leu Gly Ser Cys Ser Cys Lys
                255                 260                 265                 270

AAA GAC TCG TGA                                                              1095
Lys Asp Ser *

CTGCCCAGCGCCCCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTC  1174
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCCCTTCCTCGGGAGGCT  1253
CCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCCTGGCTACCCCCACCCTGGCTACCCCAAC  1332
GGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGC  1411
CAGGCAGCCCCGGAGGCTGGGTGGGGCTCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAATGAAACGTGAAAA  1490
AAAAAAAAAAAAGGGCGGCCGC                                                        1512
```

FIG. 1B

```
                *CnpNPCmNgGtCvNtpmYtCiCpeGYmyYtGrrC*
                  C+P PC NGG CV  +  +C+CP G   + G++C
T125pro    107  CQP-PCRNGGSCVQPG--RCRCPAG---WRGDTC    134
(EGF1)

                *CnpN..PCmNgGtCvNtp.mYtCiCpeGYm.yYtGrrC*
                  C+ +   C +  +C+NT  +Y+C+C EG++ + +G  C
T125pro    141  CSARRGGCPQ--RCINTAGSYWCQCWEGHSLSADGTLC    176
(EGF2)
```

FIG. 2

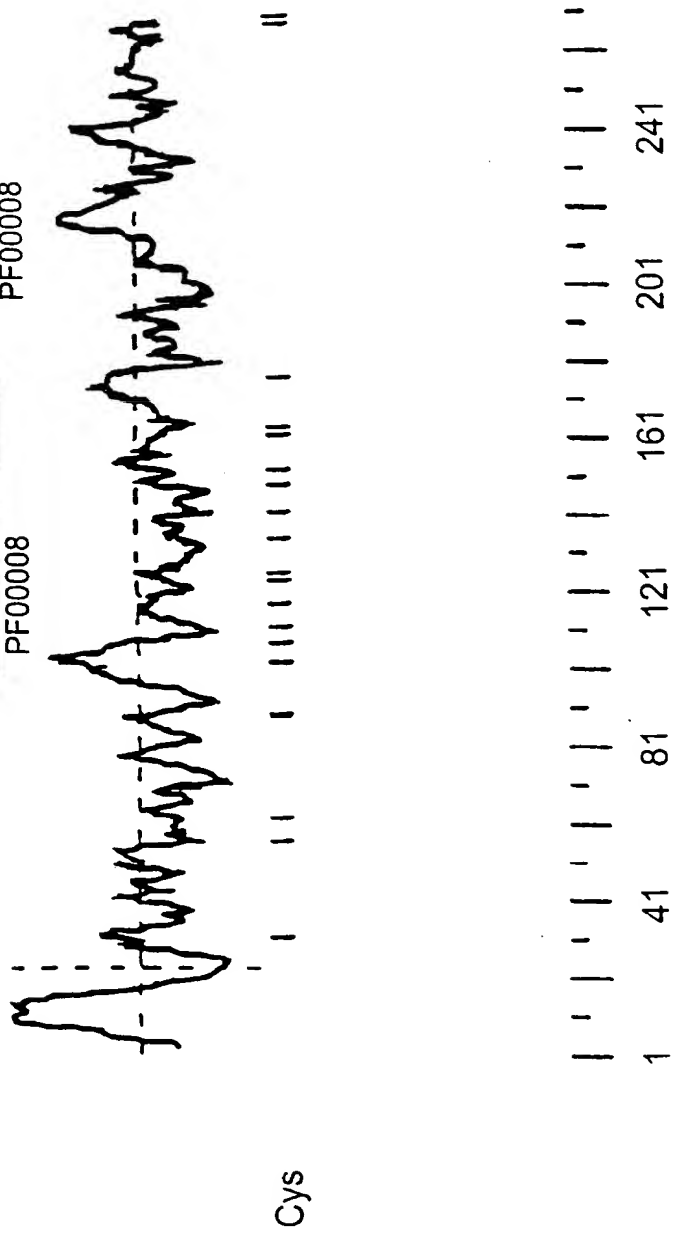


FIG. 3

```
                M  W  G  S  G  E  L  L  V  A  W  F  L  V  L  A      16
GGTACCGCCACC ATG TGG GGC TCC GGA GAA CTG CTT GTA GCA TGG TTT CTA GTG TTG GCA      60

 A  D  G  T  T  E  H  V  Y  R  P  S  R  R  V  C  T  V  G  I      36
GCA GAT GGT ACT ACT GAG CAT GTC TAC AGA CCC AGC CGT AGA GTG TGT ACT GTG GGG ATT  120

 S  G  G  S  I  S  E  T  F  V  Q  R  V  Y  Q  P  Y  L  T  T      56
TCC GGA GGT TCC ATC TCG GAG ACC TTT GTG CAG CGT GTA TAC CAG CCT TAC CTC ACC ACT  180

 C  D  G  H  R  A  C  S  T  Y  R  T  I  Y  R  T  A  Y  R  R      76
TGC GAC GGA CAC AGA GCC TGC AGC ACC TAC CGA ACC ATC TAC CGG ACT GCC TAT CGC CGT  240

 S  P  G  V  T  P  A  R  P  R  Y  A  C  C  P  G  W  K  R  T      96
AGC CCT GGG GTG ACT CCC GCA AGG CCT CGC TAT GCT TGC TGC CCT GGT TGG AAG AGG ACC  300

 S  G  L  P  G  A  C  G  A  A  I  C  Q  P  P  C  G  N  G  G     116
AGT GGG CTC CCT GGG GCT TGT GGA GCA GCA ATA TGC CAG CCT CCA TGT GGG AAT GGA GGG  360

 S  C  I  R  P  G  H  C  R  C  P  V  G  W  Q  G  D  T  C  Q     136
AGT TGC ATC CGC CCA GGA CAC TGC CGC TGC CCT GTG GGA TGG CAG GGA GAT ACT TGC CAG  420

 T  D  V  D  E  C  S  T  G  E  A  S  C  P  Q  R  C  V  N  T     156
ACA GAT GTT GAT GAA TGC AGT ACA GGA GAG GCC AGT TGT CCC CAG CGC TGT GTC AAT ACT  480

 V  G  S  Y  W  C  Q  G  W  E  G  Q  S  P  S  A  D  G  T  R     176
GTG GGA AGT TAC TGG TGC CAG GGA TGG GAG GGA CAA AGC CCA TCT GCA GAT GGG ACG CGC  540

 C  L  S  K  E  G  P  S  P  V  A  P  N  P  T  A  G  V  D  S     196
TGC CTG TCT AAG GAG GGG CCC TCC CCG GTG GCC CCA AAC CCC ACA GCA GGA GTG GAC AGC  600

 M  A  R  E  E  V  Y  R  L  Q  A  R  V  D  V  L  E  Q  K  L     216
ATG GCG AGA GAG GAG GTG TAC AGG CTG CAG GCT CGG GTT GAT GTG CTA GAA CAG AAA CTG  660

 Q  L  V  L  A  P  L  H  S  L  A  S  R  S  T  E  H  G  L  Q     236
CAG TTG GTG CTG GCC CCA CTG CAC AGC CTG GCC TCT CGG TCC ACA GAG CAT GGG CTA CAA  720

 D  P  G  S  L  L  A  H  S  F  Q  Q  L  D  R  I  D  S  L  S     256
GAT CCT GGC AGC CTG CTG GCC CAT TCC TTC CAG CAG CTG GAC CGA ATT GAT TCA CTG AGT  780

 E  Q  V  S  F  L  E  E  H  L  G  S  C  S  C  K  K  D  L  *     276
GAG CAG GTG TCC TTC TTG GAG GAA CAT CTG GGG TCC TGC TCC TGC AAA AAA GAT CTG TGA  840

CTCGAG                                                                846
```

FIG. 4

```
TTAGGATCGACCCCGCGTCTGCGGACGCGTGGGCGGACGCGTCCGCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCC  79

TGTGGACAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG  158

                                        M   A   T   P   G   L   Q   Q   H   Q   Q    11
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACC ATG GCC ACG CCT GGG CTC CAG CAG CAT CAG CAG  226

 P   P   G   P   G   R   H   R   W   P   P   P   P   G   G   A   A   P   A   P    31
CCC CCA GGA CCG GGG AGG CAC AGG TGG CCC CCA CCA CCC GGA GGA GCA GCT CCT GCC CCT  286

 V   R   G   M   T   D   S   P   P   P   G   H   P   E   E   K   A   T   P   P    51
GTC CGG GGG ATG ACT GAT TCT CCT CCG CCA GGC CAC CCA GAG GAG AAG GCC ACC CCG CCT  346

 G   G   T   G   H   E   G   L   S   G   G   A   A   D   V   A   S   G   V   G    71
GGA GGC ACA GGC CAT GAG GGG CTC TCA GGA GGT GCT GCT GAT GTG GCT TCT GGT GTT GGC  406

 S   G   R   H   R   A   R   L   P   A   R   P   *                               84
AGT GGG CGG CAC AGA GCA CGC CTA CCG GCC CGG CCG TAG                               445

GGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACC  524

TGCGACGGGCACCGGGCCTGCAGCACCTACCGGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGC  603

TGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTC  682

CCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACT  761

CTGTGTGCCCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAG  840

AGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGG  919

CACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAG  998

CGAGCAGATTTCCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGC  1077

TGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGT  1156

GACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGG  1235

GATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGT  1314

GGGCCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTG  1393

GGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGAAAAAAAAAAAAAAAAAGGGCGG  1472

CCG                                                                              1475
```

FIG. 5

```
TTAGGATCGACCCCGCGTCTGCGGACGCGTGGGCGGACGCGTCCGCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCC  79

TGTGGACAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG  158

                                        M   A   T   P   G   L   Q   Q   H   Q   Q   11
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACC ATG GCC ACG CCT GGG CTC CAG CAG CAT CAG CAG  226

 P   P   G   P   G   R   H   R   W   P   P   P   P   G   G   A   A   P   A   P   31
CCC CCA GGA CCG GGG AGG CAC AGG TGG CCC CCA CCA CCC GGA GGA GCA GCT CCT GCC CCT  286

 V   R   G   M   T   D   S   P   P   P   A   V   G   C   V   L   S   G   L   T   51
GTC CGG GGG ATG ACT GAT TCT CCT CCG CCA GCC GTA GGG TGT GTG CTG TCC GGG CTC ACG  346

 G   T   L   S   P   S   R   S   C   S   V   C   T   S   P   S   S   P   P   A   71
GGG ACC CTG TCT CCG AGT CGT TCG TGC AGC GTG TGT ACC AGC CCT TCC TCA CCA CCT GCG  406

 T   G   T   G   P   A   A   P   T   G   Q   P   P   C   R   N   G   G   S   C   91
ACG GGC ACC GGG CCT GCA GCA CCT ACC GGC CAG CCG CCA TGC CGG AAC GGA GGG AGC TGT  466

 V   Q   P   G   R   C   R   C   P   A   G   W   R   G   D   T   C   Q   S   D   111
GTC CAG CCT GGC CGC TGC CGC TGC CCT GCA GGA TGG CGG GGT GAC ACT TGC CAG TCA GAT  526

 V   D   E   C   S   A   R   R   G   G   C   P   Q   R   C   I   N   T   A   G   131
GTG GAT GAA TGC AGT GCT AGG AGG GGC GGC TGT CCC CAG CGC TGC ATC AAC ACC GCC GGC  586

 S   Y   W   C   Q   C   W   E   G   H   S   L   S   A   D   G   T   L   C   V   151
AGT TAC TGG TGC CAG TGT TGG GAG GGG CAC AGC CTG TCT GCA GAC GGT ACA CTC TGT GTG  646

 P   K   G   G   P   P   R   V   A   P   N   P   T   G   V   D   S   A   M   K   171
CCC AAG GGA GGG CCC CCC AGG GTG GCC CCC AAC CCG ACA GGA GTG GAC AGT GCA ATG AAG  706

 E   E   V   Q   R   L   Q   S   R   V   D   L   L   E   E   K   L   Q   L   V   191
GAA GAA GTG CAG AGG CTG CAG TCC AGG GTG GAC CTG CTG GAG GAG AAG CTG CAG CTG GTG  766

 L   A   P   L   H   S   L   A   S   Q   A   L   E   H   G   L   P   D   P   G   211
CTG GCC CCA CTG CAC AGC CTG GCC TCG CAG GCA CTG GAG CAT GGG CTC CCG GAC CCC GGC  826

 S   L   L   V   H   S   F   Q   Q   L   G   R   I   D   S   L   S   E   Q   I   231
AGC CTC CTG GTG CAC TCC TTC CAG CAG CTC GGC CGC ATC GAC TCC CTG AGC GAG CAG ATT  886

 S   F   L   E   E   Q   L   G   S   C   S   C   K   K   D   S   *               248
TCC TTC CTG GAG GAG CAG CTG GGG TCC TGC TCC TGC AAG AAA GAC TCG TGA               937

CTGCCCAGCGCCCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTC  1016

CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCT  1095

CCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAAC  1174

GGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGC  1253

CAGGCAGCCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGAAAA  1332

AAAAAAAAAAAAGGGCGGCCG                                                            1353
```

FIG. 6

```
TTAGGATCGACCCCGCGTCTGCGGACGCGTGGGCGGACGCGTCCGCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCC  79

TGTGGACAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG  158

                                        M   A   T   P   G   L   Q   Q   H   Q   Q    11
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACC ATG GCC ACG CCT GGG CTC CAG CAG CAT CAG CAG  226

 P   P   G   P   G   R   H   R   W   P   P   P   P   G   G   A   A   P   A   P    31
CCC CCA GGA CCG GGG AGG CAC AGG TGG CCC CCA CCA CCC GGA GGA GCA GCT CCT GCC CCT  286

 V   R   G   M   T   D   S   P   P   P   A   V   G   C   V   L   S   G   L   T    51
GTC CGG GGG ATG ACT GAT TCT CCT CCG CCA GCC GTA GGG TGT GTG CTG TCC GGG CTC ACG  346

 G   T   L   S   P   S   R   S   C   S   V   C   T   S   P   S   S   P   P   A    71
GGG ACC CTG TCT CCG AGT CGT TCG TGC AGC GTG TGT ACC AGC CCT TCC TCA CCA CCT GCG  406

 T   G   T   G   P   A   A   P   T   E   P   S   I   G   P   P   T   A   A   A    91
ACG GGC ACC GGG CCT GCA GCA CCT ACC GAA CCA TCT ATA GGA CCG CCT ACC GCC GCA GCC  466

 L   G   W   P   L   P   G   L   A   T   R   A   A   P   A   G   R   G   P   A   111
CTG GGC TGG CCC CTG CCA GGC CTC GCT ACG CGT GCT GCC CCG GCT GGA AGA GGA CCA GCG  526

 G   F   L   G   P   V   E   Q   Q   Y   A   S   R   H   A   G   T   E   G   A   131
GGC TTC CTG GGG CCT GTG GAG CAG CAA TAT GCC AGC CGC CAT GCC GGA ACG GAG GGA GCT  586

 V   S   S   L   A   A   A   A   A   L   Q   D   G   G   V   T   L   A   S   Q   151
GTG TCC AGC CTG GCC GCT GCC GCT GCC CTG CAG GAT GGC GGG GTG ACA CTT GCC AGT CAG  646

 M   W   M   N   A   V   L   G   G   A   A   V   P   S   A   A   S   T   P   P   171
ATG TGG ATG AAT GCA GTG CTA GGA GGG GCG GCT GTC CCC AGC GCT GCA TCA ACA CCG CCG  706

 A   V   T   G   A   S   V   G   R   G   T   A   C   L   Q   T   V   H   S   V   191
GCA GTT ACT GGT GCC AGT GTT GGG AGG GGC ACA GCC TGT CTG CAG ACG GTA CAC TCT GTG  766

 C   P   R   E   G   P   P   G   W   P   P   T   R   Q   E   W   T   V   Q   *   211
TGC CCA AGG GAG GGC CCC CCA GGG TGG CCC CCA ACC CGA CAG GAG TGG ACA GTG CAA TGA  826

AGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAG  905

CCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGC  984

ATCGACTCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGC  1063

CCAGCGCCCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAGA  1142

AGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCTCCCC  1221

AGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA  1300

TCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGG  1379

CAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGAAAAAAAA  1458

AAAAAAAAGGGCGGCCG                                                                  1475
```

FIG. 7

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Holtzman, Douglas

(ii) TITLE OF THE INVENTION: NOVEL MOLECULES OF THE T125-RELATED PROTEIN FAMILY AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 225 Franklin Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110-2804

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: Windows95
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 23-APR-1998

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Meiklejohn, Ph.D., Anita L.
(B) REGISTRATION NUMBER: 35,283
(C) REFERENCE/DOCKET NUMBER: 09404/048001

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617/542-5070
(B) TELEFAX: 617/542-8906
(C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1512 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: Coding Sequence
(B) LOCATION: 274...1092

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GTCGACCCAC GCGTCCGCTC AGGAGGTGCC TCCAGGCGGC CAGTGGGCCT GAGGCCCCAG     60
CAAGGGCTAG GGTCCATCTC CAGTCCCAGG ACACAGCAGC GGCCACCATG GCCACGCCTG    120
GGCTCCAGCA GCATCAGCAG CCCCCAGGAC CGGGGAGGCA CAGGTGGCCC CCACCACCCG    180
```

```
GAGGAGCAGC TCCTGCCCCT GTCCGGGGGA TGACTGATTC TCCTCCGCCA GGCCACCCAG     240
AGGAGAAGGC CACCCCGCCT GGAGGCACAG GCC ATG AGG GGC TCT CAG GAG GTG     294
                                     Met Arg Gly Ser Gln Glu Val
                                       1               5

CTG CTG ATG TGG CTT CTG GTG TTG GCA GTG GGC GGC ACA GAG CAC GCC      342
Leu Leu Met Trp Leu Leu Val Leu Ala Val Gly Gly Thr Glu His Ala
          10                  15                  20

TAC CGG CCC GGC CGT AGG GTG TGT GCT GTC CGG GCT CAC GGG GAC CCT      390
Tyr Arg Pro Gly Arg Arg Val Cys Ala Val Arg Ala His Gly Asp Pro
      25                  30                  35

GTC TCC GAG TCG TTC GTG CAG CGT GTG TAC CAG CCC TTC CTC ACC ACC      438
Val Ser Glu Ser Phe Val Gln Arg Val Tyr Gln Pro Phe Leu Thr Thr
 40                  45                  50                  55

TGC GAC GGG CAC CGG GCC TGC AGC ACC TAC CGA ACC ATC TAT AGG ACC      486
Cys Asp Gly His Arg Ala Cys Ser Thr Tyr Arg Thr Ile Tyr Arg Thr
                60                  65                  70

GCC TAC CGC CGC AGC CCT GGG CTG GCC CCT GCC AGG CCT CGC TAC GCG      534
Ala Tyr Arg Arg Ser Pro Gly Leu Ala Pro Ala Arg Pro Arg Tyr Ala
            75                  80                  85

TGC TGC CCC GGC TGG AAG AGG ACC AGC GGG CTT CCT GGG GCC TGT GGA      582
Cys Cys Pro Gly Trp Lys Arg Thr Ser Gly Leu Pro Gly Ala Cys Gly
          90                  95                  100

GCA GCA ATA TGC CAG CCG CCA TGC CGG AAC GGA GGG AGC TGT GTC CAG      630
Ala Ala Ile Cys Gln Pro Pro Cys Arg Asn Gly Gly Ser Cys Val Gln
     105                 110                 115

CCT GGC CGC TGC CGC TGC CCT GCA GGA TGG CGG GGT GAC ACT TGC CAG      678
Pro Gly Arg Cys Arg Cys Pro Ala Gly Trp Arg Gly Asp Thr Cys Gln
120                 125                 130                 135

TCA GAT GTG GAT GAA TGC AGT GCT AGG AGG GGC GGC TGT CCC CAG CGC      726
Ser Asp Val Asp Glu Cys Ser Ala Arg Arg Gly Gly Cys Pro Gln Arg
                140                 145                 150

TGC ATC AAC ACC GCC GGC AGT TAC TGG TGC CAG TGT TGG GAG GGG CAC      774
Cys Ile Asn Thr Ala Gly Ser Tyr Trp Cys Gln Cys Trp Glu Gly His
            155                 160                 165

AGC CTG TCT GCA GAC GGT ACA CTC TGT GTG CCC AAG GGA GGG CCC CCC      822
Ser Leu Ser Ala Asp Gly Thr Leu Cys Val Pro Lys Gly Gly Pro Pro
          170                 175                 180

AGG GTG GCC CCC AAC CCG ACA GGA GTG GAC AGT GCA ATG AAG GAA GAA      870
Arg Val Ala Pro Asn Pro Thr Gly Val Asp Ser Ala Met Lys Glu Glu
     185                 190                 195

GTG CAG AGG CTG CAG TCC AGG GTG GAC CTG CTG GAG GAG AAG CTG CAG      918
Val Gln Arg Leu Gln Ser Arg Val Asp Leu Leu Glu Glu Lys Leu Gln
200                 205                 210                 215

CTG GTG CTG GCC CCA CTG CAC AGC CTG GCC TCG CAG GCA CTG GAG CAT      966
Leu Val Leu Ala Pro Leu His Ser Leu Ala Ser Gln Ala Leu Glu His
                220                 225                 230

GGG CTC CCG GAC CCC GGC AGC CTC CTG GTG CAC TCC TTC CAG CAG CTC     1014
Gly Leu Pro Asp Pro Gly Ser Leu Leu Val His Ser Phe Gln Gln Leu
            235                 240                 245
```

```
GGC CGC ATC GAC TCC CTG AGC GAG CAG ATT TCC TTC CTG GAG GAG CAG      1062
Gly Arg Ile Asp Ser Leu Ser Glu Gln Ile Ser Phe Leu Glu Glu Gln
        250                 255                 260

CTG GGG TCC TGC TCC TGC AAG AAA GAC TCG TGACTGCCCA GCGCCCCAGG CTG    1115
Leu Gly Ser Cys Ser Cys Lys Lys Asp Ser
    265                 270

GACTGAGCCC CTCACGCCGC CCTGCAGCCC CCATGCCCCT GCCCAACATG CTGGGGGTCC    1175
AGAAGCCACC TCGGGGTGAC TGAGCGGAAG GCCAGGCAGG GCCTTCCTCC TCTTCCTCCT    1235
CCCCTTCCTC GGGAGGCTCC CCAGACCCTG GCATGGGATG GGCTGGGATC TTCTCTGTGA    1295
ATCCACCCCT GGCTACCCCC ACCCTGGCTA CCCCAACGGC ATCCCAAGGC CAGGTGGGCC    1355
CTCAGCTGAG GGAAGGTACG AGCTCCCTGC TGGAGCCTGG GACCCATGGC ACAGGCCAGG    1415
CAGCCCGGAG GCTGGGTGGG GCCTCAGTGG GGGCTGCTGC CTGACCCCCA GCACAATAAA    1475
AATGAAACGT GAAAAAAAAA AAAAAAAGGG CGGCCGC                             1512
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 273 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Arg Gly Ser Gln Glu Val Leu Leu Met Trp Leu Leu Val Leu Ala
  1               5                  10                  15
Val Gly Gly Thr Glu His Ala Tyr Arg Pro Gly Arg Arg Val Cys Ala
            20                  25                  30
Val Arg Ala His Gly Asp Pro Val Ser Glu Ser Phe Val Gln Arg Val
        35                  40                  45
Tyr Gln Pro Phe Leu Thr Thr Cys Asp Gly His Arg Ala Cys Ser Thr
    50                  55                  60
Tyr Arg Thr Ile Tyr Arg Thr Ala Tyr Arg Arg Ser Pro Gly Leu Ala
65                  70                  75                  80
Pro Ala Arg Pro Arg Tyr Ala Cys Cys Pro Gly Trp Lys Arg Thr Ser
                85                  90                  95
Gly Leu Pro Gly Ala Cys Gly Ala Ala Ile Cys Gln Pro Pro Cys Arg
            100                 105                 110
Asn Gly Gly Ser Cys Val Gln Pro Gly Arg Cys Arg Cys Pro Ala Gly
        115                 120                 125
Trp Arg Gly Asp Thr Cys Gln Ser Asp Val Asp Glu Cys Ser Ala Arg
    130                 135                 140
Arg Gly Gly Cys Pro Gln Arg Cys Ile Asn Thr Ala Gly Ser Tyr Trp
145                 150                 155                 160
Cys Gln Cys Trp Glu Gly His Ser Leu Ser Ala Asp Gly Thr Leu Cys
                165                 170                 175
Val Pro Lys Gly Gly Pro Pro Arg Val Ala Pro Asn Pro Thr Gly Val
            180                 185                 190
Asp Ser Ala Met Lys Glu Glu Val Gln Arg Leu Gln Ser Arg Val Asp
        195                 200                 205
Leu Leu Glu Glu Lys Leu Gln Leu Val Leu Ala Pro Leu His Ser Leu
    210                 215                 220
Ala Ser Gln Ala Leu Glu His Gly Leu Pro Asp Pro Gly Ser Leu Leu
225                 230                 235                 240
Val His Ser Phe Gln Gln Leu Gly Arg Ile Asp Ser Leu Ser Glu Gln
                245                 250                 255
Ile Ser Phe Leu Glu Glu Gln Leu Gly Ser Cys Ser Cys Lys Lys Asp
            260                 265                 270
Ser
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 819 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATGAGGGGCT CTCAGGAGGT GCTGCTGATG TGGCTTCTGG TGTTGGCAGT GGGCGGCACA     60
GAGCACGCCT ACCGGCCCGG CCGTAGGGTG TGTGCTGTCC GGGCTCACGG GGACCCTGTC    120
TCCGAGTCGT TCGTGCAGCG TGTGTACCAG CCCTTCCTCA CCACCTGCGA CGGGCACCGG    180
GCCTGCAGCA CCTACCGAAC CATCTATAGG ACCGCCTACC GCCGCAGCCC TGGGCTGGCC    240
CCTGCCAGGC CTCGCTACGC GTGCTGCCCC GGCTGGAAGA GGACCAGCGG GCTTCCTGGG    300
GCCTGTGGAG CAGCAATATG CCAGCCGCCA TGCCGGAACG GAGGGAGCTG TGTCCAGCCT    360
GGCCGCTGCC GCTGCCCTGC AGGATGGCGG GGTGACACTT GCCAGTCAGA TGTGGATGAA    420
TGCAGTGCTA GGAGGGGCGG CTGTCCCCAG CGCTGCATCA ACACCGCCGG CAGTTACTGG    480
TGCCAGTGTT GGGAGGGGCA CAGCCTGTCT GCAGACGGTA CACTCTGTGT GCCCAAGGGA    540
GGGCCCCCCA GGGTGGCCCC CAACCCGACA GGAGTGGACA GTGCAATGAA GGAAGAAGTG    600
CAGAGGCTGC AGTCCAGGGT GGACCTGCTG GAGGAGAAGC TGCAGCTGGT GCTGGCCCCA    660
CTGCACAGCC TGGCCTCGCA GGCACTGGAG CATGGGCTCC CGGACCCCGG CAGCCTCCTG    720
GTGCACTCCT TCCAGCAGCT CGGCCGCATC GACTCCCTGA GCGAGCAGAT TTCCTTCCTG    780
GAGGAGCAGC TGGGGTCCTG CTCCTGCAAG AAAGACTCG                           819
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 251 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Ala Tyr Arg Pro Gly Arg Arg Val Cys Ala Val Arg Ala His Gly Asp
  1               5                  10                  15
Pro Val Ser Glu Ser Phe Val Gln Arg Val Tyr Gln Pro Phe Leu Thr
            20                  25                  30
Thr Cys Asp Gly His Arg Ala Cys Ser Thr Tyr Arg Thr Ile Tyr Arg
        35                  40                  45
Thr Ala Tyr Arg Arg Ser Pro Gly Leu Ala Pro Ala Arg Pro Arg Tyr
    50                  55                  60
Ala Cys Cys Pro Gly Trp Lys Arg Thr Ser Gly Leu Pro Gly Ala Cys
65                  70                  75                  80
Gly Ala Ala Ile Cys Gln Pro Pro Cys Arg Asn Gly Gly Ser Cys Val
                85                  90                  95
Gln Pro Gly Arg Cys Arg Cys Pro Ala Gly Trp Arg Gly Asp Thr Cys
            100                 105                 110
Gln Ser Asp Val Asp Glu Cys Ser Ala Arg Arg Gly Gly Cys Pro Gln
        115                 120                 125
Arg Cys Ile Asn Thr Ala Gly Ser Tyr Trp Cys Gln Cys Trp Glu Gly
    130                 135                 140
His Ser Leu Ser Ala Asp Gly Thr Leu Cys Val Pro Lys Gly Gly Pro
145                 150                 155                 160
Pro Arg Val Ala Pro Asn Pro Thr Gly Val Asp Ser Ala Met Lys Glu
                165                 170                 175
Glu Val Gln Arg Leu Gln Ser Arg Val Asp Leu Leu Glu Glu Lys Leu
            180                 185                 190
```

```
Gln Leu Val Leu Ala Pro Leu His Ser Leu Ala Ser Gln Ala Leu Glu
        195                 200                 205
His Gly Leu Pro Asp Pro Gly Ser Leu Leu Val His Ser Phe Gln Gln
    210                 215                 220
Leu Gly Arg Ile Asp Ser Leu Ser Glu Gln Ile Ser Phe Leu Glu Glu
225                 230                 235                 240
Gln Leu Gly Ser Cys Ser Cys Lys Lys Asp Ser
                245                 250
```

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00, C07K 14/00 — A3

(11) International Publication Number: WO 99/54437

(43) International Publication Date: 28 October 1999 (28.10.99)

(21) International Application Number: PCT/US99/08900

(22) International Filing Date: 23 April 1999 (23.04.99)

(30) Priority Data:

09/065,363	23 April 1998 (23.04.98)	US
Not furnished	23 April 1999 (23.04.99)	US

(71) Applicant: MILLENNIUM BIOTHERAPEUTICS, INC. [US/US]; 620 Memorial Drive, Cambridge, MA 02139 (US).

(72) Inventor: HOLTZMAN, Douglas, A.; 821 Centre Street #6, Jamaica Plain, MA 01230 (US).

(74) Agent: MEIKLEJOHN, Anita, L.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110–2804 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

With international search report.

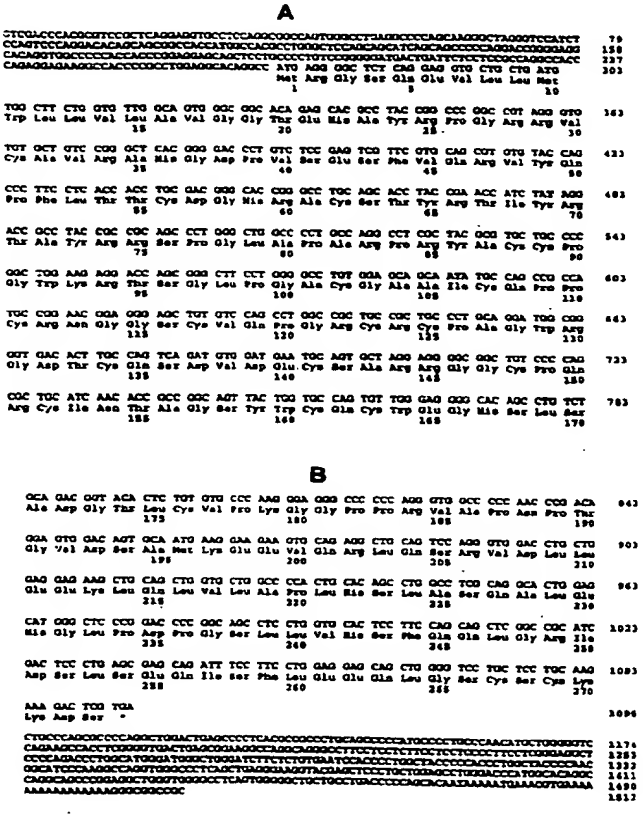
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 9 December 1999 (09.12.99)

(54) Title: NOVEL MOLECULES OF THE T125–RELATED PROTEIN FAMILY AND USES THEREOF

(57) Abstract

Novel T125 polypeptides, proteins, and nucleic acid molecules are disclosed. In addition to isolated, full–length T125 proteins, the invention further provides isolated T125 fusion proteins, antigenic peptides and anti–T125 antibodies. The invention also provides T125 nucleic acid molecules, recombinant expression vectors containing a nucleic acid molecule of the invention, host cells into which the expression vectors have been introduced and non–human transgenic animals in which a T125 gene has been introduced or disrupted. Diagnostic, screening and therapeutic methods utilizing compositions of the invention are also provided.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/08900

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/00; C07K 14/00
US CL :435/69.1, 325, 363; 536/23.5; 530/350
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 325, 363; 536/23.5; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE
search terms: T125 protein, Tango protein, EGF-like domain

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 98/57983 A2 (ZYMOGENETICS, INC.) 23 December 1998, pages 45-48, 17-18 and claims 1 and 6	1-8, 10, 12
X,P	Database GENBANK on STN, NCI-CGAP. 'National Cancer Institute, Cancer Genome Anatomy Project, Tumor Gene Index'. AN AI146849, 26 October 1998.	1-3
X --- Y	Database GENBANK on STN. HILLIER et al. 'zx07b12.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone'. AN AA448958, 04 June 1997.	1, 3-5, 12 --- 6,7

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
24 SEPTEMBER 1999

Date of mailing of the international search report
20 OCT 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
ELIANE LAZAR-WESLEY
JOYCE BRIDGERS
PARALEGAL SPECIALIST
CHEMICAL MATRIX

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/08900

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10 and 12-13

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/08900

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-10 and 12-13, drawn to an isolated nucleic acid molecule of SEQ ID No:1 or 3, an isolated polypeptide of SEQ ID No:2 or 4, a host cell and a method of making a polypeptide .

Group II, claim 11, drawn to an antibody.

Group III, claims 14-16 and 20-21, drawn to a method of detecting the presence of a polypeptide in a sample by using a compound that binds to the polypeptide, and a method for identifying a compound.

Group IV, claims 17-19, drawn to a method of detecting the presence of a nucleic acid.

Group V, claims 22-23, drawn to a method for modulating the activity of the polypeptide.

The inventions listed as Groups I-V do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical feature of Group I is the nucleic acid of SEQ ID No:1 or 3, the polypeptide of SEQ ID no:2 or 4, and a method of using the nucleic acid for making a polypeptide. Group II is directed to a different product. Groups III-V are directed to different methods that use different products and steps that do not correspond to the special technical feature of Group I.